Title	Some Distribution-Free Multivariate Comparison Procedures (統計的多変量解析の研究報告集 II)
Author(s)	田村, 亮二
Citation	数理解析研究所講究録 (1968), 60: 47-56
Issue Date	1968-12
URL	http://hdl.handle.net/2433/107845
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

数理解析研究所講究録
60巻 1968年 47-56

Some distribution-free multivariate comparison procedures

九州芸工大　田村 亮二

§1. 序

c個の p-変量処理母集団 π_i $(i=1,\cdots,c)$ と対照母集団 π_0 の分布関数をそれぞれ $F_i(\underset{\sim}{x})=F(\underset{\sim}{x}-\underset{\sim}{\theta}_i)$, $F_0(\underset{\sim}{x})=F(\underset{\sim}{x})$ とする。$F(\underset{\sim}{x})$ の連続性(または絶対連続性)は仮定するがその関数形は未知である。昨年のシンポジュームで次の多重比較[1]を考察した。与えられた定数ベクトル $\underset{\sim}{a}'=(a^{(1)},\cdots,a^{(p)})$, $\forall a^{(h)}\geqq 0$ に対して, $\underset{\sim}{a}'\underset{\sim}{\theta}_i=\Delta_i$ とし

(1)
$\Delta_i>0$ のとき π_i は π_0 より better
$\Delta_i\leqq 0$ のとき π_i は π_0 より not better

という基準が定められているとき

(2)　$Pr[\,\forall \underset{\sim}{\theta}_i=\underset{\sim}{0}$ のとき π_0 が best として select$\,]=1-\alpha$

を満たし, π_0 より better なものと not better なものに分離せよ。筆者はこれに対して Rank procedure を提唱し, 正規分布 $F(\underset{\sim}{x})$ の仮定で用いられる procedure との比較を行った。

今回は (1) の拡張である次の基準の下の問題を考察する。

与えられた q 個の定数ベクトル $a_h' = (a_h^{(1)}, \cdots, a_h^{(p)})$, $\forall a_h^{(k)} \geqq 0$

$h=1,\cdots,q$ に対して, $\Delta_i' = (\Delta_i^{(1)}, \cdots, \Delta_i^{(q)})$, $\Delta_i^{(h)} = a_h' \theta_i$ とし,

$$\begin{array}{lll} & \Delta_i \leqq 0 \ (\Delta_i = 0 \text{ も含む}),\ i=1,\cdots,c \text{ のとき} & \pi_0 \text{ は best} \\ (3) & \Delta_i \geqq 0 \ (\Delta_i = 0 \text{ は含まず}) \text{ のとき} & \pi_i \text{ は } \pi_0 \text{ より better} \\ & \Delta_i \not\geqq 0 \ (\Delta_i \geqq 0 \text{ の否定}) \text{ のとき} & \pi_i \text{ は } \pi_0 \text{ より not better} \end{array}$$

という基準が定められているとき

(4) $P_r[\forall \Delta_i = 0$ のとき π_0 が best として select$] \geqq 1-\alpha$

を満たし π_0 より better なものと not better なものに分離する procedure 如何.

さらに $q=1$ のときの結果と前回の結果との比較を試みる.

(昨年のシンポジュームで山本教授より示唆)

§2. 定義と補助定理

補助定理 1. p-変量の確率変数 X の cdf $F(x-\theta)$, a_h $(h=1,\cdots,q)$ は与えられ p-ベクトルで $A'=[a_1,\cdots,a_q]$ の rank は q とする。そのとき $Y^{(h)} = a_h' X$, $h=1,\cdots q$ の同時確率密度は $g(y-\Delta)$ の形で与えられる。ただし $\Delta^{(h)} = a_h' \theta$, $\Delta = (\Delta^{(1)}, \cdots, \Delta^{(q)})$.

証明は初等的であるから略. なお $F(x)$ が共分散行列 Σ をもてば $g(y)$ の共分散行列は $A\Sigma A'$ になることも明らか.

この補助定理から我々の問題は次のように形式化 きる．

q-変量の処理母集団 π_i，その分布関係 $G_i(\underset{\sim}{y}) = G(\underset{\sim}{y} - \underset{\sim}{\Delta}_i)$．$(i=1,\cdots,c)$．対照 π_0 の分布関数 $G_0(\underset{\sim}{y}) = G(\underset{\sim}{y})$ で $G(\underset{\sim}{y})$ の連続性(または絶対連続性)は仮定するがその関数形は未知．処理の良さについての基準は(3)で定められている．今 G_j からの標本 $\{\underset{\sim}{Y}_{j1},\cdots,Y_{jn_j}\}$，$\underset{\sim}{Y}_{j\alpha} = (Y_{j\alpha}^{(1)},\cdots,Y_{j\alpha}^{(q)})$，$\alpha=1,\cdots,n_j$，$j=0,1,\cdots,c$，$\sum_{j=0}^{c} n_j = N$ に基いて(4)を満たし π_0 より better なものと not better なものに分離する方法を求めること．簡単のため

D_0 ： 対照 π_0 が best であるという判定

$D_{i_1\cdots i_r}$ ： π_{i_β} $(\beta=1,\cdots,r)$ は π_0 より better で残りの π_{j_γ} $(\gamma=1,\cdots,s,\ r+s=c)$ は π_0 より not better という判定

とする．

定義 1．

$$(5)\qquad n_j T_{Nj}^{(h)} = \sum_{\alpha=1}^{n_j} Z(R(Y_{j\alpha}^{(h)})),\quad j=0,1,\cdots c,\quad h=1,\cdots,q,$$

$R(Y_{j\alpha}^{(h)})$ = 才 h 成分全体(個数は N)の中での $Y_{j\alpha}^{(h)}$ の rank

$Z(1) < \cdots < Z(N)$ ： $\Phi(x)$ (標準正規分布 $N(0,1)$) からの大きさ N の順序統計量

$$(6)\qquad \begin{aligned} \underset{\sim}{T}_N^{(h)\prime} &= (\hat{T}_{N1}^{(h)},\cdots,\hat{T}_{Nc}^{(h)}) & h&=1,\cdots,q \\ \underset{\sim}{T}_{Ni}^{\prime} &= (\hat{T}_{Ni}^{(1)},\cdots,\hat{T}_{Ni}^{(q)}) & i&=1,\cdots,c \end{aligned}$$

$$\hat{T}_{Ni}^{(h)} = \left(\frac{n_0 n_i}{n_0+n_i}\right)^{\frac{1}{2}} (T_{Ni}^{(h)} - T_{N0}^{(h)}) \ .$$

定義２．

(7) $$n_j \bar{Y}_{Nj}^{(h)} = \sum_{\alpha=1}^{n_j} Y_{j\alpha}^{(h)} , \qquad j=0,1,\cdots,c \ , \quad h=1,\cdots,q$$

(8) $$\bar{\underset{\sim}{Y}}_N^{(h)\prime} = (\hat{\bar{Y}}_{N1}^{(h)}, \cdots, \hat{\bar{Y}}_{Nc}^{(h)}) \qquad h=1,\cdots,q$$

$$\bar{\underset{\sim}{Y}}_{Ni}^{\prime} = (\hat{\bar{Y}}_{Ni}^{(1)}, \cdots, \hat{\bar{Y}}_{Ni}^{(q)}) \qquad i=1,\cdots,c$$

$$\hat{\bar{Y}}_{Ni}^{(h)} = \left(\frac{n_0 n_i}{n_0+n_i}\right)^{\frac{1}{2}} (\bar{Y}_{Ni}^{(h)} - \bar{Y}_{N0}^{(h)}) / (\underset{\sim}{a}_h' S \underset{\sim}{a}_h)^{\frac{1}{2}}$$

S は Σ の一致推定量

補助定理２． $\forall \underset{\sim}{\Delta}_i = \underset{\sim}{0}$ のとき $\underset{\sim}{T}_N^{(h)}$ の分布は exact に c-変量正規分布 $N(\underset{\sim}{0}, \Lambda)$ である．$\bar{\underset{\sim}{Y}}_N^{(h)}$ は漸近的に $N(\underset{\sim}{0}, \Lambda)$

$$\Lambda = [\lambda_{ij}] \ , \quad \lambda_{ij} = \begin{cases} 1 & i=j \\ [\lambda_i \lambda_j / (\lambda_0+\lambda_i)(\lambda_0+\lambda_j)]^{\frac{1}{2}} & i \neq j \end{cases}$$

$\lambda_j = n_j / N$ ．

証明． $\underset{\sim}{T}_N^{(h)}$ については Bell-Doksum [1] から．また $\bar{\underset{\sim}{Y}}_N^{(h)}$ については中心極限定理と Mann-Wald [2] から．

補助定理３． $\underset{\sim}{\Delta}_i = \underset{\sim}{\delta}_i / \sqrt{N}$ ， $\underset{\sim}{\delta}_i = (\delta_i^{(1)}, \cdots, \delta_i^{(q)})$ のとき

(i) $\underset{\sim}{T}_N^{(1)}, \cdots, \underset{\sim}{T}_N^{(q)}$ の同時分布は漸近的に $N(\mu, \Lambda \otimes \Gamma)$

(9) $$\mu_i^{(h)} = \left(\frac{\lambda_0 \lambda_i}{\lambda_0+\lambda_i}\right)^{\frac{1}{2}} \delta_i^{(h)} \int_{-\infty}^{\infty} \frac{d}{dy} \Phi^{-1}(G^{(h)}(y)) \, dG^{(h)}(y) \ , \quad \mu = [\mu_i^{(h)}]$$

(10) $$\gamma_{hk} = \begin{cases} 1 & h=k \\ \iint \Phi^{-1}(G^{(h)}(x)) \Phi(G^{(k)}(y)) \, dG^{(h,k)}(x,y) & h \neq k \end{cases}$$

$$\Gamma = [\gamma_{hk}]$$

$G^{(h)}$, $G^{(h,k)}$は$G(\underset{\sim}{y})$の第h成分，第(h,k)成分の周辺分布.

(ii) $\underset{\sim}{\bar{Y}}_N^{(1)}, \cdots, \underset{\sim}{\bar{Y}}_N^{(g)}$ の同時分布は漸近的に $N(\nu, \Lambda\otimes\Omega)$

(11) $\nu_i^{(h)} = \left(\frac{\lambda_0\lambda_i}{\lambda_0+\lambda_i}\right)^{\frac{1}{2}}\delta_i^{(h)} \Big/ (\underset{\sim}{a}_h'\Sigma\underset{\sim}{a}_h)^{\frac{1}{2}}$, $\nu = [\nu_i^{(h)}]$

(12) $$\omega_{hk} = \begin{cases} 1 & h=k \\ \underset{\sim}{a}_h'\Sigma\underset{\sim}{a}_k / (\underset{\sim}{a}_h'\Sigma\underset{\sim}{a}_h)^{\frac{1}{2}}(\underset{\sim}{a}_k'\Sigma\underset{\sim}{a}_k)^{\frac{1}{2}} & h\neq k \end{cases}$$

$\Omega = [\omega_{hk}]$

証明. $n_j\,\delta_{Nj}^{(h)} = \sum_{\alpha=1}^{n_j} \mathcal{E}[Z(R(Y_{j\alpha}^{(h)}))]$.

$\underset{\sim}{\delta}_N^{(h)\prime} = (\hat{\delta}_{N1}^{(h)}, \cdots, \hat{\delta}_{Nc}^{(h)})$, $\hat{\delta}_{Ni}^{(h)} = \left(\frac{n_0 n_i}{n_0+n_i}\right)^{\frac{1}{2}}(\delta_{Ni}^{(h)} - \delta_{N0}^{(h)})$

とおくと，$\sqrt{N}(T_{Nj}^{(h)} - J_{Nj}^{(h)}) \xrightarrow{(P)} 0$ $(N\to\infty)$ が Bell-Doksum [1] によって示されている．かくて $\sqrt{N}(\underset{\sim}{T}_N^{(1)}, \cdots, \underset{\sim}{T}_N^{(g)})$ と $\sqrt{N}(\underset{\sim}{\delta}_N^{(1)}, \cdots, \underset{\sim}{\delta}_N^{(g)})$ とは漸近的に同じ分布に従う．さらに後者が漸近的に $N(\mu, \Lambda\otimes\Gamma)$ に従うことは田村[3] の結果である．重の代りにある条件を満たす H に対しても同様の結果が得られる．

§3. Selection procedures.

Procedure N.

(13) $\underset{\sim}{T}_{Ni} \leq \underset{\sim}{z}_\alpha$, $i=1,\cdots,c$ ならば D_0 を採択せよ

$\underset{\sim}{T}_{Ni_\beta} > \underset{\sim}{z}_\alpha$, $\beta=1,\cdots,r$, $\underset{\sim}{T}_{Nj_\gamma} \not> \underset{\sim}{z}_\alpha$, $\gamma=1,\cdots,s$, $r+s=c$ ならば $D_{i_1\cdots i_r}$ を採択せよ

(14) $$\int_{-\infty}^{z_\alpha}\cdots\int_{-\infty}^{z_\alpha} n(\underset{\sim}{0}, \Lambda)\, d\underset{\sim}{y} = 1-\frac{\alpha}{g} \quad , \quad \underset{\sim}{z}_\alpha' = (z_\alpha, \cdots, z_\alpha)$$

Procedure M

(15) $\bar{\underset{\sim}{Y}}_{Ni} \leqq \underset{\sim}{z}_\alpha$, $i=1,\cdots,c$ ならば D_0 を採択せよ

$\bar{\underset{\sim}{Y}}_{Ni_\beta} > \underset{\sim}{z}_\alpha$, $\beta=1,\cdots,r$, $\bar{\underset{\sim}{Y}}_{Nj_\gamma} \not> \underset{\sim}{z}_\alpha$, $\gamma=1,\cdots,s$, $r+s=c$

ならば $D_{i_1\cdots i_r}$ を採択せよ.

定理 1. Procedure N では (4) は"strictly"に成立するが Procedure M では漸近的にである.

証明. $P_r[\,\forall \underset{\sim}{\Delta}_i = \underset{\sim}{0}$ のとき π_0 が best として select $]$

$= P_r[\ \underset{\sim}{T}_{Ni} \leqq \underset{\sim}{z}_\alpha,\ i=1,\cdots,c \mid \forall \underset{\sim}{\Delta}_i = \underset{\sim}{0}\ ]$

$= P_r[\ \underset{\sim}{T}_N^{(h)} \leqq \underset{\sim}{z}_\alpha,\ h=1,\cdots,q \mid \forall \underset{\sim}{\Delta}_i = \underset{\sim}{0}\ ]$

Bonferroni の不等式から

$\geqq\ 1 - \sum_{h=1}^{q} P_r[\ \underset{\sim}{T}_N^{(h)} \not\leqq \underset{\sim}{z}_\alpha \mid \forall \underset{\sim}{\Delta}_i = \underset{\sim}{0}\ ]$

$=\ 1-q + \sum_{h=1}^{q} P_r[\ \underset{\sim}{T}_N^{(h)} \leqq \underset{\sim}{z}_\alpha \mid \forall \underset{\sim}{\Delta}_i = \underset{\sim}{0}\ ]$

$=\ 1-\alpha$ ((14) を用いた)

Procedure M では $P_r[\underset{\sim}{T}_N^{(h)} \leqq \underset{\sim}{z}_\alpha \mid \forall \underset{\sim}{\Delta}_i = \underset{\sim}{0}] \sim 1-\frac{\alpha}{q}$ であるため (4) は strictly にではなく漸近的に成立.

定理 2. 各 Procedure で次式が漸近的に成立する.

(16) $P_r[\,\forall \underset{\sim}{\Delta}_i \leqq \underset{\sim}{0}$ のとき π_0 が best として select $]$

$\geqq\ 1-\alpha$

証明. $\underset{\sim}{\Delta}_i = \underset{\sim}{\delta}_i/\sqrt{N}$, $\underset{\sim}{\delta}_i \leqq \underset{\sim}{0}$ のとき証明すればよい.
定理 1 のときと同様にして,

(16)の左辺 $= \Pr[\forall T_{Ni} \leqq z_\alpha \mid \forall \Delta_i = \delta_i/\sqrt{N},\ \delta_i \leqq 0]$

$$\geqq (1-q) + \sum_{h=1}^{q} \Pr[T_N^{(h)} \leqq z_\alpha \mid \forall \Delta_i = \delta_i/\sqrt{N},\ \delta_i \leqq 0]$$

補助定理3より

$$\sim (1-q) + q \int_{-\infty}^{z_\alpha} \cdots \int_{-\infty}^{z_\alpha} n(\mu^{(h)}, \Lambda)\, dy$$

$$\mu^{(h)} = (\mu_1^{(h)}, \cdots, \mu_c^{(h)}) \leqq 0$$

$$\geqq (1-q) + q \int_{-\infty}^{z_\alpha} \cdots \int_{-\infty}^{z_\alpha} n(0, \Lambda)\, dy$$

$$= 1-\alpha$$

系1. $G^{(h,k)}(x,y)$ が平均ベクトル 0 (一般性を失わず), 共分散行列 $\begin{pmatrix} a_h' \Sigma a_h & a_h' \Sigma a_k \\ a_k' \Sigma a_h & a_k' \Sigma a_k \end{pmatrix}$ の正規分布のとき, 2つのProcedureは漸近的に同等である.

証明. 上の仮定の下では, (9) (10) から容易に

$$\mu_i^{(h)} = \left(\frac{\lambda_0\lambda_i}{\lambda_0+\lambda_i}\right)^{\frac{1}{2}} \delta_i^{(h)} \Big/ (a_h' \Sigma a_h)^{\frac{1}{2}}$$

$$\gamma_{hk} = a_h' \Sigma a_k \Big/ (a_h' \Sigma a_h)^{\frac{1}{2}} (a_k' \Sigma a_k)^{\frac{1}{2}} \qquad h \neq k$$

となり, $T_N^{(1)}, \cdots, T_N^{(q)}$ の漸近分布は $\bar{T}_N^{(1)}, \cdots, \bar{T}_N^{(q)}$ のそれと同じくなる.

§4. $q=1$ のとき.

$q=1$ の場合は勿論上述の特別な場合であるが, もう少し詳しく論ずることができる. またこの結果と前回の結果[4]の比

較も考察する．π_jからの標本を$\{Y_{j1},\cdots,Y_{jn_j}\}$とする．

定義3．

(17) $$n_j T_{Nj}(H) = \sum_{\alpha=1}^{n_j} Z(R(Y_{j\alpha})), \quad j=0,1,\cdots,c$$

$Z(1)<\cdots<Z(N)$ ： 既知の分布関数 $H(z)$ からの順序統計量

（§2 では $H(z)=\Phi(z)$ のときのみ問題にした）

$$\hat{T}_{Ni}(H) = \left(\frac{n_0 n_i}{n_0+n_i}\right)^{\frac{1}{2}}\left(T_{Ni}(H)-T_{N0}(H)\right)/\sigma$$

$$\sigma^2 = \int_0^1 H^{-1}(t)^2\,dt - \left(\int_0^1 H^{-1}(t)\,dt\right)^2$$

Procedure H

(18) $\hat{T}_{Ni}(H) \leq z_\alpha$, $i=1,\cdots,c$ ならば D_0 を採択せよ

$\hat{T}_{Ni_\beta}(H) > z_\alpha$, $\beta=1,\cdots,r$, $\hat{T}_{Nj_\gamma}(H) \leq z_\alpha$, $\gamma=1,\cdots,s$, $r+s=c$

ならば $D_{i_1\cdots i_r}$ を採択せよ

ただし $$\int_{-\infty}^{z_\alpha}\cdots\int_{-\infty}^{z_\alpha} n(\underset{\sim}{0},\Lambda)\,d\underset{\sim}{y} = 1-\alpha .$$

$H(z)=\Phi(z)$, $H(z)=z$ にとったときの Procedure をそれぞれ Procedure H_N, H_U で表わす．

系2． Procedure H_N は strictly に次式を満足する． $H(z)$ が Bell-Doksum の条件[1] を満せば Procedure H は漸近的に(19)を満足する．

(19) $P_r[\,\forall\Delta_i=0$ のとき π_0 が best として select$\,] = 1-\alpha$

証明． Procedure H_N に対しては補助定理2と定理1とが

ら．一般のHに対しては[1]と定理2から．

定理3．　$\Delta_i=\delta_i/\sqrt{N}$ のとき，$D_{i_1\cdots i_r}$ が正しい確率は漸近的に次式で与えられる．

(20)
$$P_{i_1\cdots i_r}(H)\sim\int_{z_\alpha-\mu_{i_1}}^{\infty}\cdots\int_{z_\alpha-\mu_{i_r}}^{\infty}\int_{-\infty}^{z_\alpha-\mu_{j_1}}\cdots\int_{-\infty}^{z_\alpha-\mu_{j_s}}n(\underset{\sim}{0},\Lambda)\,d\underset{\sim}{y}$$

(21)
$$\mu_i=\left(\frac{\lambda_0\lambda_i}{\lambda_0+\lambda_i}\right)^{\frac{1}{2}}\delta_i\int_{-\infty}^{\infty}\frac{d}{dx}H^{-1}(G(x))\,dG(x)\Big/\sigma$$

証明．　$P_{i_1\cdots i_r}(H)=P[\hat{T}_{Ni_\beta}>z_\alpha,\ \beta=1,\cdots,r,\ \hat{T}_{Nj_\gamma}\leqq z_\alpha$

$\gamma=1,\cdots,s,\quad r+s=c\ |\ \Delta_i=\delta_i/\sqrt{N},\ \delta_{i_\beta}>0,\ \beta=1,\cdots,r$

$\delta_{j_\gamma}\leqq 0,\ \gamma=1,\cdots,s\ ]$

補助定理3を用いて容易に上の結果が得られる．

定理4．前回の Procedure W [4] の Procedure H に対する漸近相対効率 $e_{W,H}$ は

(22)
$$e_{W,H}=\left[\sigma\int_{-\infty}^{\infty}\frac{d}{dx}J[F(x)]\,dF(x)\Big/(\underset{\sim}{a}'\Pi\,\underset{\sim}{a})^{\frac{1}{2}}\int_{-\infty}^{\infty}\frac{d}{dy}H^{-1}(G(y))\,dG(y)\right]^2$$

で与えられる．

証明．(20) (21) と前稿の結果を比較すれば容易．

Procedure W と H の対応するものを比較すれば

(a) Normal score type

$J(t)=\Phi^{-1}(t)$, $H(z)=\Phi(z)$ の場合で $e_{W,H_N}=1$

(b) Wilcoxon type

$J(t)=t$, $H(z)=z$ の場合で，そのときは (22)から

(23) $$e_{W,H_v} = (\underset{\sim}{a}'\Sigma\underset{\sim}{a})/(\underset{\sim}{a}'B\underset{\sim}{a})$$

$$B=[b_{hk}], \qquad b_{hk} = \begin{cases} 1 & h=k \\ \frac{6}{\pi}\sin^{-1}\frac{\rho_{hk}}{2} & h\neq k \end{cases}$$

$$\Sigma = [\rho_{hk}] \qquad \rho_{hh}=1$$

さて (23) は

$$e_{W,H_v} = (\underset{\sim}{a}'\Sigma\underset{\sim}{a})\Big/\Big[\underset{\sim}{a}'\Sigma\underset{\sim}{a} - \sum_{h\neq k} a^{(h)}a^{(k)}\Big\{\rho_{hk}-\frac{6}{\pi}\sin^{-1}\frac{\rho_{hk}}{2}\Big\}\Big]$$

かくて $\forall\rho_{hk}\geqq 0$ のとき $\rho_{hk}-\frac{6}{\pi}\sin^{-1}\frac{\rho_{hk}}{2}\geqq 0$ となり $e_{W,H_v}\geqq 1$ を得る。$\forall\rho_{hk}\leqq 0$ のときは同様に $e_{W,H_v}\leqq 1$

文　献

[1] Bell-Doksum : Some new distribution-free statistics. A.M.S. 36 (1965) 203-214

[2] Mann-Wald : On stochastic limit and order relationships. A.M.S. 14 (1943)

[3] Tamura : Multivariate nonparametric several-sample tests. A.M.S. 37 (1966)

[4] 田村亮二 : Some nonparametric methods for multivariate analysis. 数理解析研講究録 44 (1968)